APPLICATION OF HIGHER ORDER CUMULANTS FOR SEISMIC WAVELET RECONSTRUCTION

Yajun Li[1]*, Baojun Yang[1], and Yue Li[2]

*[*1] College of Geoexploration Science and Technology, Jilin University, 6 Ximinzhu Street, Changchun 130026, China*
Email: lyj_2872@sohu.com
[2] College of Communication Engineering, Jilin University, 5372 Nanhu Road, Changchun 130012, China

ABSTRACT

According to the fact that high order cumulants (HOC) retain the phase information of signals and the HOC of the Gaussian color noise is always equal to zero, a new method of wavelet reconstruction is provided in this paper, based on 4th-order cumulants of non-Gaussian seismic signals. The feasibility of this method is demonstrated by the simulation of wavelet estimation for synthetic seismic traces. Furthermore, the seismic wavelet of field data processed with this method can be reconstructed correctly.

Keywords: Signal processing, Higher order cumulants (HOC), Seismic wavelet, Seismic exploration

1 INTRODUCTION

In seismology exploration, it is very important that seismic wavelets be estimated accurately. This estimation is fundamental for seismic traces deconvolution, migration, feature extraction, and geophysical interpretation (Sheridd & Geldart, 1995).

Statistical estimation is very valuable for seismic wavelet estimation and deconvolution. The common method of deconvolution concerns 2nd-order statistics (SOS) of the data on the assumption that the noise is white and the wavelet is considered the minimum phase. In the past 20 years, high order statistical (HOS) theories and methods have become one of the most important fields in signal processing, for HOS can be used to extract the non-minimum phase wavelet from Gaussian color noise (Damien & Salah, 2005; Lazear, 1993; Mauricio, Sacchi & Ulrych, 2000).

Because the background noise of seismic signals tends to be Gaussian, the HOC of noise will diminish as the sampling size increases. However, as the order k increases, the variance of a kth-order sample estimator increases also. Thus, it is preferable to keep k as low as possible. Based on actual measured data, seismic traces are generally regarded as having a symmetric non-Gaussian distribution (Giannakis & Tsatsanis, 1994). Hence, when 3rd-order cumulants of the data vanish, 4th-order cumulants include significant phase information which 2nd-order statistics do not have. Therefore, this paper puts forward a feasible method of analysis to enhance seismic signal detection and classification using 4th-order cumulants.

Assuming that $b(1) = 1$, Xianda's (1999) cumulants slice method is used to pick up MA parameters $b(n)$ ($n = 1, 2, \cdots, N$). This assumption is non-applicable for the seismic wavelet. This paper will derive matrix equations based on 4th-order cumulants when $b(1)$ is a discretionary value not equal to zero. These equations have many uses. Theoretically, this method can remove Gaussian white/color noise effectively. Simulation of synthetic data and application of field data also demonstrate that it is a new promising method for seismic signal processing.

2 SEISMIC SIGNAL MODEL

It is well known that a received seismic trace can be described (Sheridd & Geldart, 1995) by

$$y(n) = \sum_i w(i)r(n-i) + v(n)\,, \tag{1}$$

where $y(n)$ is the seismic trace with noise, $w(n)$ is the seismic wavelet with arbitrary phase, and $v(n)$ is the additive Gaussian white/color noise. The earth reflectivity sequence $r(n)$ is assumed to be zero-mean and i.i.d (independently identically distributed). Moreover, $r(n)$ is generally regarded as having a symmetric non-Gaussian distribution and it is independent with $v(n)$.
Our goal is to reconstruct the wavelet $w(n)$ by using the finite data record $\{y(n)\}$ ($n = 1, 2, \cdots, N$).

3 MATRIX EQUATIONS BASED ON FOURTH-ORDER CUMULANTS

For computational convenience without loss of generality, one can choose *i* from 1 to *lw* in Equation 1, where *lw* is the length of the seismic wavelet. Thus, Equation 1 can be written as

$$y(n)=\sum_{i=1}^{lw} w(i)r(n-i)+v(n)\,. \tag{2}$$

Obviously, Equation 2 corresponds to the MA model whereas $w(1) \neq 1$.

Using the BBR (Bartlett-Brillinger-Rosenblatt) formula (Xianda, 1999), the 4th-order cumulants of Equation 2 can be calculated as follows

$$C_{4y}(\tau_1,\tau_2,\tau_3)=C_{4r}(\tau_1,\tau_2,\tau_3)\sum_{i=1}^{lw} w(i)w(i+\tau_1)w(i+\tau_2)w(i+\tau_3)+C_{4v}(\tau_1,\tau_2,\tau_3)\,, \tag{3}$$

where τ_1, τ_2 and τ_3 are time delays.

For the *k*th-order ($k \geqslant 3$) cumulants of additive Gaussian processes to vanish, when *n* is large enough, the 4th-order cumulants of $r(n)$ can be written as follows

$$C_{4r}(\tau_1,\tau_2,\tau_3)=r_4\delta(i-\tau_1,i-\tau_2,i-\tau_3)\,, \tag{4}$$

where r_4 is kurtosis of the reflectivity sequence, $\delta(\tau_1, \tau_2, \tau_3)$ is multidimensional impulse function.

Thus, Equation 3 can be transformed to the following form

$$C_{4y}(\tau_1,\tau_2,\tau_3)=r_4\sum_{i=1}^{lw} w(i)w(i+\tau_1)w(i+\tau_2)w(i+\tau_3)\,, \tag{5}$$

where $w(i) = 0$, when $i < 1$ or $i > lw$.

Substituting $\tau_1 = \tau$, $\tau_2 = \tau_3 = 0$ in Equation 5, one obtains the one-dimensional (1-D) slice of $C_{4y}(\tau_1, \tau_2, \tau_3)$ as follows

$$C_{4y}(\tau,0)=C_{4y}(\tau,0,0)=r_4\sum_{i=1}^{lw} w^3(i)w(i+\tau)\,. \tag{6}$$

Similarly, only substituting $\tau_3 = 0$ in Equation 5, one obtains the 2-D slice of $C_{4y}(\tau_1,\tau_2,\tau_3)$

$$C_{4y}(\tau_1,\tau_2)=C_{4y}(\tau_1,\tau_2,0)=r_4\sum_{i=1}^{lw} w^2(i)w(i+\tau_1)w(i+\tau_2)\,. \tag{7}$$

Substituting $\tau = lw$-1 in Equation 6 and $\tau_1 = lw$-1, $\tau_2 = n$ in Equation 7, after transformation one has

$$w(n+1)=\frac{C_{4y}(lw-1,n)w(1)}{C_{4y}(lw-1,0)}\,. \tag{8}$$

Letting $w(i+\tau) = b(i+\tau)$ and making a variable substitution in Equation 6 can yield

$$C_{4y}(\tau,0)=r_4\sum_{i=1}^{lw} w^3(i)b(i+\tau)=r_4\sum_{j=1}^{lw} w^3(j-\tau)b(j)=r_4\sum_{i=1}^{lw} w^3(i-\tau)b(i)\,. \tag{9}$$

Then substituting Equation 8 into Equation 9, one gets

$$r_4\sum_{i=1}^{lw} b(i)C_{4y}^3(lw-1,i-\tau-1)w^3(1)=C_{4y}(\tau,0)C_{4y}^3(lw-1,0)\,. \tag{10}$$

Similarly, letting $w(i) = b(i)$ and substituting Equation 8 into Equation 6, one obtains

$$r_4\sum_{i=1}^{lw} b^3(i)C_{4y}(lw-1,i+\tau-1)w(1)=C_{4y}(\tau,0)C_{4y}(lw-1,0)\,. \tag{11}$$

With $\tau = lw$-1 and $b(1) = w(1)$, one obtains

$$r_4=\frac{C_{4y}^2(lw-1,0)}{b^4(1)C_{4y}(lw-1,lw-1)}\,. \tag{12}$$

Substituting Equation 12 into Equation 10 yields

$$\sum_{i=1}^{lw}\tfrac{b(i)}{b(1)}C_{4y}^3(lw-1,i-\tau-1)=C_{4y}(\tau,0)C_{4y}(lw-1,0)C_{4y}(lw-1,lw-1)\,. \tag{13}$$

Now, one obtains the linear equations based on 4th-order cumulants concerning parameters $b(i)$. They form the matrix equation

$$\boldsymbol{A}_y\boldsymbol{B}=\boldsymbol{C}_y, \tag{14}$$

where

$$
\boldsymbol{A}_y = \begin{bmatrix}
C_{4y}^3(lw-1,lw-1) & 0 & \cdots & 0 & 0 \\
C_{4y}^3(lw-1,lw-2) & C_{4y}^3(lw-1,lw-1) & \cdots & 0 & 0 \\
\vdots & \vdots & \ddots & \vdots & \vdots \\
C_{4y}^3(lw-1,1) & C_{4y}^3(lw-1,2) & \cdots & C_{4y}^3(lw-1,lw-1) & 0 \\
C_{4y}^3(lw-1,0) & C_{4y}^3(lw-1,1) & \cdots & C_{4y}^3(lw-1,lw-2) & C_{4y}^3(lw-1,lw-1) \\
0 & C_{4y}^3(lw-1,0) & \cdots & C_{4y}^3(lw-1,lw-3) & C_{4y}^3(lw-1,lw-2) \\
\vdots & \vdots & \ddots & \vdots & \vdots \\
0 & 0 & \cdots & 0 & C_{4y}^3(lw-1,0)
\end{bmatrix},
$$

$$
\boldsymbol{B} = \begin{bmatrix} b(1) & b(2) & \cdots & b(lw) \end{bmatrix}^T / b(1),
$$

$$
\boldsymbol{C}_y = \begin{bmatrix}
C_{4y}(-(lw-1),0)C_{4y}(lw-1,0)C_{4y}(lw-1,lw-1) \\
C_{4y}(-(lw-2),0)C_{4y}(lw-1,0)C_{4y}(lw-1,lw-1) \\
\vdots \\
C_{4y}(-1,0)C_{4y}(lw-1,0)C_{4y}(lw-1,lw-1) \\
C_{4y}(0,0)C_{4y}(lw-1,0)C_{4y}(lw-1,lw-1) \\
C_{4y}(1,0)C_{4y}(lw-1,0)C_{4y}(lw-1,lw-1) \\
\vdots \\
C_{4y}(lw-1,0)C_{4y}(lw-1,0)C_{4y}(lw-1,lw-1)
\end{bmatrix}.
$$

Solving this matrix equation using generalized inverses, one can obtain the value of $b(1)$, $b(2)$, $\cdots$, $b(lw)$, where the denominator $b(1)$ only influences the amplitudes of the assessed wavelet not its waveform.

4 FOURTH-ORDER CUMULANTS OF OBSERVATION DATA

High-order cumulants do not vary with a shift in the mean. Hence, it is convenient to define them under the assumption of zero mean. If the process has a non-zero mean, we subtract the mean and then apply the following definitions to the resulting process.

4th-order cumulants are functions of observed seismic traces $y(n)$, which assumes a zero-mean and stationary random process. Thus, 4th-order cumulants can be defined approximately (Xianda, 1999):

$$
C_{4y}(\tau_1,\tau_2,\tau_3) = \frac{1}{k}\sum_{n=1}^{k} y(n)y(n+\tau_1)y(n+\tau_2)y(n+\tau_3) - C_{2y}(\tau_1)C_{2y}(\tau_3-\tau_2) - C_{2y}(\tau_2)C_{2y}(\tau_3-\tau_1) - C_{2y}(\tau_3)C_{2y}(\tau_2-\tau_1) . \tag{15}
$$

The auto-correlation function $C_{2y}(\tau)$ can be obtained from $\{ y(n)\}$ ($n = 1, 2, \cdots , N$)

$$
C_{2y}(\tau) = \frac{1}{k}\sum_{n=1}^{k} y(n)y(n+\tau) , \tag{16}
$$

where $C_{2y}(-\tau) = C_{2y}(\tau)$.

5 SIMULATIONS

In order to test the feasibility of the cumulants method in seismic exploration, two simulations are given in this section. One uses synthetic seismic data and the other uses field seismic data.

5.1 Simulation of synthetic seismic data

5.1.1 Reconstruction of zero-phase Ricker wavelet

The synthetic data are generated by a convolutional model shown in Figure 1. The Ricker wavelet, which has a zero phase, dominant frequency of 40(Hz) with sample interval of 4(ms), is adopted. The reflectivity sequence is assumed to be a series of random numbers from -0.5 to 0.5 and has a symmetric non-Gaussian distribution. Then the zero-mean Gaussian colored noise is added.

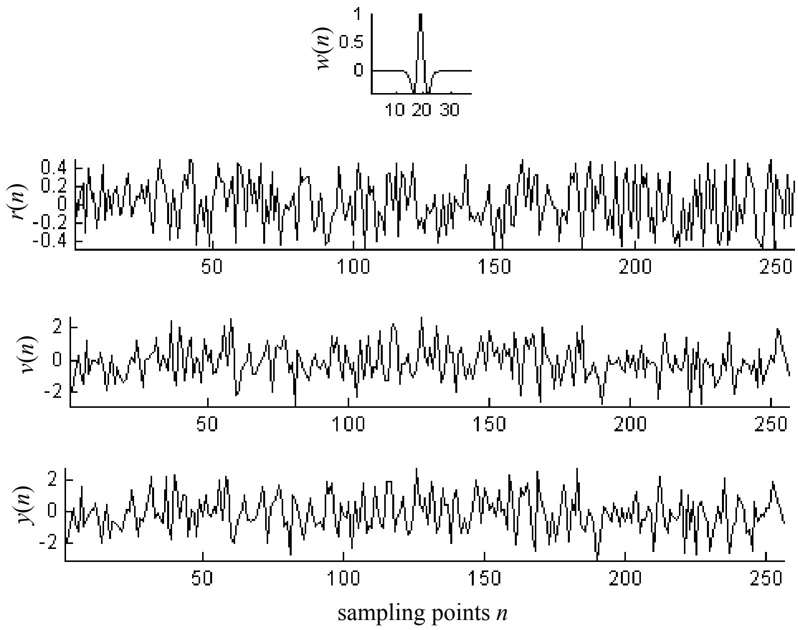


Figure 1. Synthesization of a seismogram. $w(n)$, Ricker wavelet; $r(n)$, earth reflectivity sequence; $v(n)$, additive Gaussian color noise; $y(n)$, synthetic seismogram.

After Equation 14 has been computed, the Ricker wavelet and the reconstructed seismic wavelet are plotted in Figure 2. It is shown that these two waveforms are coincident in the rough.

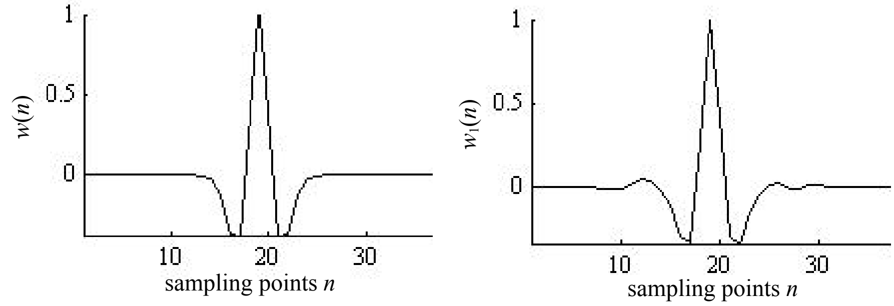


Figure 2. Wavelet reconstruction for seismogram synthesized by zero-phase Ricker wavelet. $w(n)$, Ricker wavelet(left); $w_1(n)$, reconstructed wavelet by (14) (right).

5.1.2 Reconstruction of a mixed-phase wavelet

The constructed mixed-phase seismic wavelet, which has a dominant frequency of 40(Hz) and sample interval of 4(ms), is constructed. The reflectivity sequence and the noise are the same as in the above example.

The constructed and estimated seismic wavelets are plotted in Figure 3. It is shown that there are some magnitude differences and slight waveform distortions between the constructed wavelet and the reconstructed one, but these differences do not distort the seismic deconvolution results badly. The result indicates that a seismic signal can be detected by the proposed matrix equations method based on 4th-order cumulants.

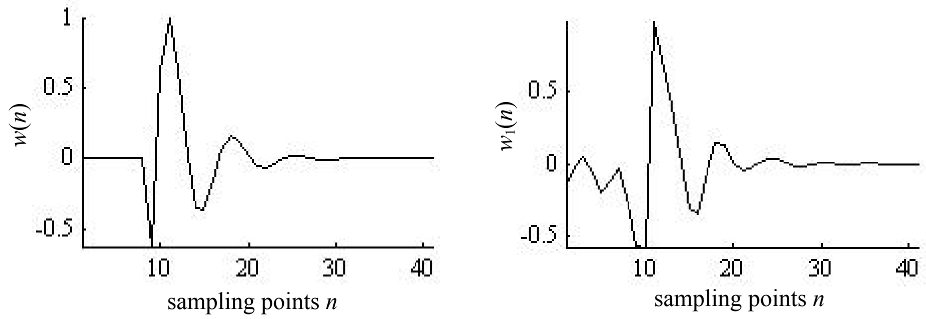
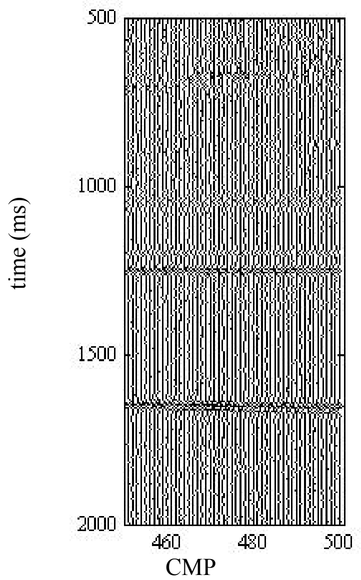


Figure 3. Wavelet reconstruction for seismogram synthesized by mixed-phase wavelet. $w(n)$, mixed-phase wavelet(left); $w_1(n)$, reconstructed wavelet by (14) (right).

5.2 Application to field data

This paper presents an application of the method mentioned above to a post-stack seismic time section from the Daqing oil field (Heilongjiang, China). This data set is obtained from a number of groups of detectors. On each sensor point, the seismic signal lasts 5(s) with a sampling interval of 1(ms), i.e., there are a total of 5000 sampling points.

From the large numbers of seismic data sets, a set of 50 CMPs is continuously selected for this application and partly shown in Figure 4. The vertical axis is marked with the time denoting the depth of the stratum. The horizontal axis is marked with the seismic trace numbers within the CMP gather.



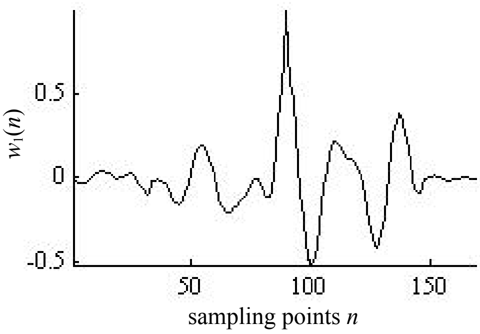


Figure 5. Wavelet reconstruction for actual seismogram. w_1(n), reconstructed wavelet by (14)

Figure 4. Actual seismic time section

Using the matrix equations method based on 4th-order cumulants, the reconstructed seismic wavelet is shown in Figure 5. It is shown that real seismic wavelet is mixed-phase. One can estimate the dominant frequency of the seismic wavelet as 45(Hz), which is in accord with the real seismic wavelet.

6 CONCLUSIONS

In this paper, it is shown that new matrix equations based on 4th-order cumulants are sufficient for the identification of generally non-minimum phase wavelets $w(n)$ ($n = 1, 2, \dots, N$), where $w(1) \neq 1$. The method does not follow the traditional assumption that the wavelet must be of a minimum phase and is valid for wavelets of any phase. These equations have more extensive practicality. It can be said that the high order cumulants technique has considerable potential for seismic data processing.

7 ACKNOWLEDGEMENTS

This work was supported by the National Natural Science Foundation of China (Grant Nos. 40374045, 40574051). The authors appreciate the valuable suggestions from Professor Shuxun Wang and Professor Shixue Dong.

8 REFERENCES

Damien, M. & Salah, B. (2005) Multiway filtering based on fourth-order cumulants. *EURASIP Journal on Applied Signal Processing (7)*, pp.1147-1158

Giannakis, G.B. & Tsatsanis, M. (1994) Time-domain tests for Gaussianity and time-reversibility. *IEEE Transactions on Signal Processing 42(12)*, pp.3460-3472

Lazear, G.D. (1993) Mixed-phase wavelet estimation using fourth-order cumulants. *Geophysics 58(7)*, pp.1042-1051

Mauricio, D., Sacchi, T., & Ulrych, J. (2000) Nonminimum-phase wavelet estimation using higher order statistics. *Society of Exploration Geophysicists 19(1)*, pp.80-83

Sheridd, C.H. & Geldart, L.P. (1995) *Exploration seismology* (second edition), Cambridge, Britain: Cambridge University Press

Xianda, Z. (1999) *Time Series Analysis: Higher Order Statistics Method*, Beijing, China: TsingHua University Publishing House